Roteiro 14 de Setembro de 2012

1. Abrir ***glass.arff***
2. Clicar em *edit* e analisar a base:
    1. Quantos atributos têm a base?
    2. Existe algum atributo com valor ausente?
    3. Qual o número de classes?
    4. Você conseguiu entender o que cada atributo representa?
3. Acesse a pagina da UCI *(*http://archive.ics.uci.edu/ml/) e descubra mais informações relevantes sobre o conjunto.
4. Colocar no modo de teste ***percentage split*** com 70%
5. Aplicar ***5-nn (lazy.iBk)***
6. Anotar a acurácia.

7. Abrir ***glass.arff***
8. Aplicar filtro ***unsupervised.attribute.RandomSubset[numAttributes=0.1]***
9. Anotar os atributos que restaram, inclusive o atributo classe
10. Colocar no modo de teste ***percentage split*** com 70%
11. Aplicar ***5-nn***
12. Anotar a acurácia.

13. Abrir ***glass.arff***
14. Aplicar filtros ***supervised.attribute.attributeSelection[evaluator = WrapperSubsetEval[classifier=5-nn]]***
15. Anotar os atributos que restaram, inclusive o atributo classe
16. Colocar no modo de teste ***percentage split*** com 70%
17. Aplicar ***5-nn***
18. Anotar a acurácia.

19. Abrir ***glass.arff***
20. Remover manualmente atributos que foram anotados no item 8 e 14
21. Anotar os atributos que restaram, inclusive o atributo classe
22. Colocar no modo de teste ***percentage split*** com 70%
23. Aplicar ***5-nn***
24. Anotar a acurácia.

25. A acurácia melhorou ou piorou quando se usaram menos atributos? Qual técnica de seleção de atributos conseguiu maior acurácia? Você acha que para a base de dados ***glass*** faz sentido utilizar seleção de atributos? Por quê?
26. Repita o procedimento para as bases: ***iris*** e ***diabates***